Plan Overview

A Data Management Plan created using DMPonline

Title: FAIR-CURES-RO

Creator: Shoaib Sufi

Principal Investigator: Stian Soiland-Reyes, Carole Goble

Data Manager: Stian Soiland-Reyes

Affiliation: Software Sustainability Institute

Template: University of Manchester Generic Template

ORCID iD: 0000-0003-1219-2137

Project abstract:

The purpose of this project is to assist Elsevier in their NIH Data Commons Pilot https://commonfund.nih.gov/commons award “FAIR data to drive CURES” https://www.elsevier.com/about/press-releases/science-and-technology/elsevier-and-seven-bridges-receive-nih-data-commons-grant-for-biomedical-data-analysis , in particular to consult on the implementation of Research Object http://www.researchobject.org/ support in Customer’s existing Mendeley Data platform https://data.mendeley.com/ and Seven Bridges Fair4CURES platform, both interacting with the proposed NIH Data Commons Research Object Composer and KC2 GUID Broker. In addition University of Manchester will use its position to assist this effort in aligning and collaborating with other ongoing community standard efforts, for improved interoperability and sustainability, and to help position the NIH Data Commons outcomes within the larger archiving and reproducibility community.

ID: 31495

Last modified: 05-10-2018

Grant number / URL: N/A

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

FAIR-CURES-RO

Manchester Data Management Outline

1. Is this project already funded?

- Yes

Will you be applying for funding from any of the following sources? If your funder isn't listed, please enter in the free text box provided.

- Wellcome Trust

Elsevier, subcontractor to NIH Data Commons FAIR4CURES project.

3. Is The University of Manchester the lead institution for this project?

- No (please provide details of the lead institution below and your role in the project)

The lead is Elsevier, subcontractor to NIH Data Commons funded FAIR4CURES project.

4. What data will you use in this project (please select all that apply)?

- Not acquire or re-use data (please provide details)

This project will only generate source code and not acquire or generate data. Data used for examples/testing are from NIH Data Commons and Seven Bridges platform.

5. Where will the data be stored and backed-up during the project lifetime?

- Other storage system (please list below)

Open Source code maintained at Elsevier's GitLab installation, to be published to https://github.com/researchobject with releases assigned DOI from https://zenodo.org/

6. If you will be using Research Data Storage, how much storage will you require?

- Not applicable

7. If you have a contractual agreement with a 3rd party data provider will any of the data associated with this project be sourced from, processed or stored outside of the institutions and groups stated on your agreement?

- Not applicable

8. How long do you intend to keep your data for after the end of your project (in years)?

- > 20 years

Zenodo provides long term storage of software archives.

Questions about personal information

Personal information or personal data, the two terms are often used interchangeably, relates to identifiable living individuals. Special category personal data is more sensitive information such as medical records, ethnic background, religious beliefs, political opinions, sexual orientation and criminal convictions or offences information. If you are not using personal data then you can skip the rest of this section.

Please note that in line with data protection law (the General Data Protection Regulation and Data Protection Act 2018), personal information should only be stored in an identifiable form for as long as is necessary for the project; it should be pseudonymised (partially de-identified) and/or anonymised (completely de—identified) as soon as practically possible. You must obtain the appropriate ethical approval in order to use identifiable personal data.

- No sensitive or personal data

10. Please provide details of how you plan to store, protect and ensure confidentiality of the participants' information as stated in the question above.

n/a

11. If you are storing personal information will you need to keep it beyond the end of the project?

- Not applicable

12. Sharing person identifiable information can present risks to participants' privacy, researchers and the institution. Will the participants' information (personal and/or sensitive) be shared with or accessed by anyone outside of the University of Manchester? This includes using 3rd party service providers such as cloud storage providers or survey platforms.

- No

13. If you will be sharing personal information outside of the University of Manchester will the individual or organisation you are sharing with be outside the EEA?

- Not applicable

14. Are you planning to use the personal information for future purposes such as research?

- No

15. Who will act as the data custodian or information asset owner for this study?

Stian Soiland-Reyes

16. Please provide the date on which this plan was last reviewed (dd/mm/yyyy).

05/10/2018

Project details

What is the purpose of your research project?

Develop a web microservice for creating and validating Research Objects for use within NIH Data Commons and other platforms.

What policies and guidelines on data management, data sharing, and data security are relevant to your research project?

n/a

Responsibilities and Resources

Who will be responsible for data management?

Stian Soiland-Reyes

What resources will you require to deliver your plan?

Zenodo, available for free with no cost.

Data Collection

What data will you collect or create?

No data will be collected or created

How will the data be collected or created?

n/a

Documentation and Metadata

What documentation and metadata will accompany the data?

n/a

Ethics and Legal Compliance

How will you manage any ethical issues?

While this project do not handle data, the software developed here may be used to handle with personal data.
However the deployed microservice will be hosted by Elsevier, not University of Manchester. The developed service will not be storing data, only prepare it for storage in other NIH Data Commons providers like Mendeley Data.

Extra care will be taken so the service does not unintentially leak private or sensitive information, such as by using authenticated SSL-encrypted communication.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Elsevier will own copyright of software produced, but it will be licensed open source under the Apache License 2.0 so that the University (or anyone else) may freely distribute, reuse or further develop the code.

Storage and backup

How will the data be stored and backed up?

Data is only transient at microservice and do not need to be backed up.

How will you manage access and security?

The microservice will use SSL transport (https) and unique session keys or URIs to avoid unintentional cross-communicatoin between users. Users of the microservice will be at freedom to specify the third-party service where data should ultimately be stored, e.g. Mendeley Data.

Selection and Preservation

Which data should be retained, shared, and/or preserved?

None

What is the long-term preservation plan for the dataset?

No datasets

Data Sharing

How will you share the data?

No data will be made or shared.

The software produced will be open source and shared publicly through GitHub and Zenodo DOIs through regular releases.

Are any restrictions on data sharing required?

No